THE MILBANK QUARTERLY
A MULTIDISCIPLINARY JOURNAL OF POPULATION HEALTH AND HEALTH POLICY

New Evidence on the Allocation of NIH Funds across Diseases

BHAVEN N. SAMPAT,[1] KRISTIN BUTERBAUGH,[2] and MARCEL PERL[2]

[1]*Mailman School of Public Health, Columbia University;* [2]*Mt. Sinai School of Medicine*

Context: The responsiveness of NIH (National Institutes of Health) funding to disease burden is a long-standing issue of policy interest. Previous analyses of this issue have been hindered by data constraints, have not specified channels through which the NIH funding process could be responsive to disease considerations, and have not examined differences across NIH institutes and centers.

Methods: We collected data from the NIH's new RCDC (Research, Condition, and Disease Categorization) database on funding for 107 diseases in 2008 and linked these to data on deaths and hospitalizations for these diseases. We used RCDC data and information from another NIH database—RePORTER—to determine institute-specific funding for these diseases and also funding by award type. We used these data to examine the overall responsiveness of NIH funding to disease burden, within-institute responsiveness, and the responsiveness of different types of NIH awards.

Findings: Overall, we found a strong and statistically significant relationship between NIH funding and deaths and hospitalizations associated with a disease. We detected some evidence that more "applied" grant mechanisms—in particular, funding for clinical trials—are more responsive than other types of funding. We also found evidence of differences across institutes in their extent of responsiveness.

Conclusions: Overall, the data suggest that NIH funding is responsive to the two measures of disease burden. More applied grant mechanisms also may serve as "safety valves" in the allocation process, allowing Congress, disease advocacy groups, and others to apply pressure to address particular health priorities in a

Address correspondence to: Bhaven N. Sampat, 600 West 168th Street, 6th Floor, New York, NY 10032 (email: bns3@columbia.edu).

The Milbank Quarterly, Vol. 91, No. 1, 2013 (pp. 163–185)

more fine-grained way than is possible through investigator-initiated "basic" research grants alone.

Keywords: National Institutes of Health (U.S.), resource allocation.

The NIH (National Institutes of Health) is the largest single funder of biomedical research in the world. The agency's cross-disease allocation decisions reveal a fundamental trade-off in health policy: in a world of finite resources, who shall live? (Fuchs 1975). Not surprisingly, the fundamental question of whether NIH's cross-disease allocation patterns are appropriate—that is, whether particular diseases command too much or too little funding—has been a perennial topic of debate throughout the agency's history (Rettig 1977; Strickland 1972). A memorable recent episode in this saga occurred in the early 1990s, with congressional concerns that AIDS, breast cancer, and other high-profile diseases got more than they should, given their burden (Johnson 1998).

Theoretical perspectives on the "optimal" allocation of health resources also suggest that disease burden ought to be a consideration in NIH funding choices (Lichtenberg 2001; Weisbrod 1983; Zeckhauser 1967). Previous empirical work on the NIH's allocation process (Gillum et al. 2011; Gross, Anderson, and Powe 1999; Lichtenberg 2001; Mushkin 1979) found mixed evidence on NIH's responsiveness to standard measures of disease burden. But this research faced a major data constraint, a lack of consistently collected data on funding by disease. One indicator of this, in his response to one pioneering study (Gross, Anderson, and Powe 1999), the then NIH director Harold Varmus noted that "the method of coding research dollars, while consistent from year to year for a specific disorder, may differ from one disease to another" (Varmus 1999b, 1914). Similarly, in congressional testimony on priority setting, Varmus observed: "Coding of funds by disease category across the NIH, though useful for some purposes, is also inherently imprecise" (Varmus 1997). An influential IOM report (IOM 1998, 5) also raised serious concerns about the measurement of NIH funding, as well as the NIH leadership's unwillingness to stand behind them, noting that the data that then existed were "not of the quality that they should be and [that] NIH should work to improve the data and better explain the data to the public."

In 2008, the NIH created the Research, Condition, and Disease Categorization (RCDC) system, which uses standard methodologies to classify funds by area. This was in part a response to the IOM report's call for "a regular process [of disease categorization] that is supported by a clear philosophy and that is sensitive to [a] scientific understanding of the meaning of such data and their interpretation" (IOM 1998, 40) and also to the 2006 NIH Reauthorization Act, which required "the Director of NIH, when reporting on research activities relating to a specific disease, disorder, or other adverse health condition," to "present information in a standardized format."

Whereas each of the NIH's twenty-seven institutes and centers had previously linked its grants to diseases in an ad hoc, nonstandard, way, the RCDC instituted new, standard, category definitions to classify grants, developed with input from disease groups and the scientific community. Thus under the old system, the agency had routinely provided disease-specific funding figures tentatively and with many caveats, whereas after implementation of the RCDC, the agency's stance shifted. The NIH website now states that "RCDC provides consistent and transparent information to the public about NIH-funded research. For the first time, a complete list of all NIH-funded projects related to each category is available" (NIH 2011a).

We used the RCDC data to examine the responsiveness of NIH funding to two measures of disease burden, U.S. deaths and hospitalizations. In addition to more reliable funding figures than were available before, the RCDC has other advantages. For example, while earlier work was based on a smaller sample of diseases (typically fewer than thirty), the RCDC provides funding data for 215 categories (about half of which are diseases). One of the contributions of this article is that we used the new data to reassess the relationships between funding and measures of disease burden.

A second contribution of our article is that we explored channels through which the agency's funding choices could respond to disease burden. Previous analyses relating funding to disease burden said nothing about the mechanisms through which the NIH might respond to disease burden. These analyses simply assumed that such mechanisms exist, even though this is not obvious. Numerous observers have noted that the NIH peer review process, which focuses mainly on scientific considerations, may lack the ability to respond to disease burden in a more fine-tuned way (IOM 1998; Varmus 1997). The 1998 IOM report

on NIH priority setting (discussed earlier) suggested: "There is a sense that NIH has evolved mechanisms for judging scientific opportunity and merit that surpass its capabilities for assessing and being influenced by public health needs" (IOM 1998, 8). Although the NIH (1997) acknowledges an obligation to consider "the incidence, severity, and cost of specific disorders" (4) and claims data on these are "closely monitored" (Varmus 1999a, 1914), it is unclear where in the allocation process these variables are used to prioritize investments (Callahan 1999; IOM 1998).

The IOM report also observed that while "basic" research funded through investigator-initiated grants are the agency's mainstay, "other mechanisms, such as grants for research centers, clinical trials, and R&D contracts, tend to be more directed in nature and more tied to an institute's mission. They are usually solicited by NIH through requests for applications (RFAs) rather than initiated by extramural scientists. . . . They also tend to be used more in problem-oriented research efforts" (IOM 1998, 19).

These more applied mechanisms have been crucial to several major initiatives at the NIH (e.g., the War on Cancer, the Artificial Kidney Program, the NIH Artificial Heart Program), reflecting the difficulties of targeting research through investigator-initiated peer-reviewed grants (Sampat 2012). They therefore may serve as "safety valves" in the allocation process, enabling the agency to respond from time to time to political pressures for more targeted funding (Sampat 2012). The RCDC database also can disaggregate funds by disease to individual grants, thereby allowing an examination of the degrees to which different types of grants—those emanating from RFAs, clinical trials grants, center grants, and contracts—are responsive to the disease burden.

A third contribution of our article is to provide new evidence on allocation choices within individual NIH institutes. Some scholars have suggested that one way in which patterns of NIH funding could respond to disease burden is through congressional allocations to the NIH's twenty-seven institutes and centers (Mowery, Nelson, and Martin 2010). Some of the institutes (e.g., the National Cancer Institute) specialize in particular diseases, and a common political strategy for increasing funding for specific diseases has been through the creation of new institutes (National Research Council 2003; Strickland 1972). Various observers have also proposed that linkages between disease and institute are fuzzy (Mushkin 1979; NIH 1997), which would make cross-institute allocations a blunt priority-setting mechanism (see also

Sampat 2012). The data allow, too, an estimation of the extent of each of the institutes' disease-based priority setting. This is useful not only to compare them with the responsiveness of NIH funding overall but also to address the suggestion made by previous observers that different NIH institutes may have differing commitments to priority setting (IOM 1998). Kastor (2010) emphasized the institutes' autonomy and the considerable power and discretion of individual institute directors. Indeed, the institutes have been characterized as "mega-silos" (Kastor 2010, 41), and the NIH as "a group of largely independent fiefdoms . . . institutes led by directors who consider themselves the final authority on matters affecting their programs" (Cohen 1993, 1675). In this context, the various institutes' responsiveness to disease burden seems plausible.

Next we discuss the funding and disease burden data, and then we report our analyses relating NIH funding to disease burden. Finally, we summarize our findings and limitations and discuss their implications.

Data

RCDC Categories and Diseases

We began with the RCDC table provided by the NIH and focused on funding in 2008. Sociologists and historians of science have persuasively argued that there is no such thing as objective disease categorization schemes, that even the International Classification of Disease categories were developed for historically contingent and sometimes political reasons (Bowker 1996; Bowker and Starr 1999). The same is true of the RCDC categories, which build on the existing infrastructure of "diseases of interest" developed "to enable NIH to respond to requests from Congress or others" (IOM 1998, 21). Historically, the NIH tracked its funding internally (albeit imperfectly, for the reasons discussed earlier) for more than two hundred categories but reported annually to Congress on only about fifty of these categories. The RCDC includes data for the full set of categories, 213 in 2008. We noted earlier that the RCDC represents what the NIH considers the state of the art in funding by disease and a major advance over previous funding figures that were calculated inconsistently across institutes and over time. However, reflecting difficulties common to any disease classification (Bowker and Starr 1999), and perhaps exacerbated when classifying research grants (whose effects

may spill over—serendipitously or otherwise—across disease areas or may be difficult to relate to any specific disease at all; see Varmus 1997, 1999a, 1999b), the categories remain imperfect. Nonetheless, in the following analyses, we regarded the RCDC categories and figures as given, since they are currently the best available data.

We began by determining which of these 213 RCDC categories mapped to actual diseases. To do so, we mapped each of the RCDC categories to the ninth (ICD-9) and tenth (ICD-10) editions of the International Classification of Diseases (ICD). We did not attempt to link risk factors ("smoking"), broader causes ("climate change"), or fields of science ("Biotechnology," "Basic Research") to ICD-9 or ICD-10 codes, instead focusing on only those categories directly linked to specific diseases. We linked the categories to ICD codes based on names and erred toward omitting categories in which the name alone did not permit a reliable link. Of the 213 conditions on the RCDC list, we were able to link 109 to a disease. A supplementary appendix (appendix S1) discusses coding rules in detail, and a supplementary table (table S1) shows the RCDC category, funding, and ICD codes for these RCDC categories.

The categories in the RCDC are not mutually exclusive; that is, the same grant can be associated with numerous categories. This reflects two realities. The first is that research can span multiple categories, so that some grants could be classified as addressing both breast cancer and pancreatic cancer, and so the funds from these grants would be counted in each category. Rather than attempt to apportion these grants across diseases, in relating funding to measure of disease burden, we assumed (as the NIH does in its aggregate calculations of funds by disease for its RCDC funding table) that the full amount of each grant was associated with each of the disease categories in which it is categorized. (The main results are robust to alternative weighing schemes, however.) A more problematic issue is that the categories are not nested, so that grants for breast cancer count for both "Breast Cancer" and "Cancer." This is potentially more troubling for estimation purposes, since if the propensity to include both broad categories and subcategories varies with burden of diseases, the resulting estimates of NIH responsiveness to disease burden could be misleading. Accordingly, we flagged as "supercategories" the twelve categories that subsumed 95 percent of the grants from one or more other categories on the RCDC and treated them separately in the analyses.

NIH Funding by Mechanism and Institute

The RCDC allows the disaggregation of funds to individual grants in each category; that is, all the grants in a category are listed, together with funding amounts and grant numbers. The grant numbers indicate grant mechanisms (e.g., research projects, clinical trials, contracts, centers) and other information.

Overall, 56,417 grants mapped to the RCDC in 2008. The average number of categories per grant was 4.8; the median was 4. About two-thirds of these grants (40,378) mapped to at least one of the 109 disease categories in the RCDC (as opposed to the other categories, e.g., for fields of science or risk factors). On average, each of these 37,236 grants mapped to 1.9 disease categories, and the median number of disease categories per grant was 2.

We counted as clinical trial grants any that mapped to one or more of the diseases and were also categorized as "Clinical Trials" in the RCDC. Funding for all clinical and nonclinical grants in a given category was summed to create disease-specific clinical and nonclinical research stocks. We defined center grants as those with the activity codes associated with centers as denoted by a recent IOM publication on extramural research centers: P30, P50, P60, U54, P20, M01, P40, U42, P41, U41, P51, and G12 (IOM 2004, 160–62). Funding for all center and noncenter grants in a given category was summed to create disease-specific center and noncenter research stocks. We defined research contracts as awards with activity codes beginning with "N" (in practice, N01, N02, N03, N43, and N44 awards).

Next, we determined the funding institute and center for each grant. Since grants and contracts can be funded by multiple institutes (about 5 percent are), we used information from the RCDC field "funding institute/center" rather than the grant number (which lists only the primary funding institute) to make this determination. Based on this information, we also constructed funding for each RCDC disease by institute, which we then used in the within-institute analyses of allocation patterns.

We also linked the RCDC grants to information from another NIH database, RePORTER (NIH 2011b), which contains information on all NIH grants and contracts funded in 2008, not just those linked to an RCDC category. Overall, the grants and contracts in RePORTER totaled $26 billion in funding in 2008. Of all NIH grants and contracts, 83 percent (accounting for 94 percent of all funding) were associated with

an RCDC category. About 55 percent of NIH grants and contracts (accounting for 67 percent of all NIH funding) were associated with one of the 109 categories in the RCDC that could be linked to a specific disease.

Linking the RCDC and RePORTER data also facilitated our undertaking a separate examination of grants that were initiated by investigator, versus grants that took the form of more targeted RFAs. To carry this out, we collected information on all 2008 grants that emanated from RFAs, as indicated in the RePORTER data. Funding for all RFA and non-RFA grants in a given category was summed to create disease-specific RFA and non-RFA funding stocks.

Overall, a minority of the NIH awards were for the more "applied" mechanisms. Thus the average share (across the 109 categories) of funding associated with clinical trials is 18 percent; the average share associated with contracts, 4 percent; and centers, 7 percent. On average, 14 percent of the funding was for awards emanating from RFAs. (These categories are not mutually exclusive, e.g., contracts resulting from RFAs can be used to fund clinical trials.)

Disease Burden

The correct measure of disease burden has been contested (Gold and Fryback 2002). For our analyses, we used two measures for which reliable data are available across a large number of diseases. The first is the number of deaths associated with a disease. We obtained information on the average number of deaths for each disease in 2005 and 2006 from the Multiple Cause of Death Mortality files (from the National Vital Statistics System of the National Center for Health Statistics). The Vital Statistics data record a census of all deaths by ICD-10 codes. The 2005/2006 period allows for a lagged response by the NIH. We averaged the data over two years, since in most of the analyses, we took natural logs of the disease burden measures. Aggregating two years of data limits the number of diseases when deaths are zero and natural logs are undefined.

The second measure is hospital stays. For this measure, we obtained estimates of hospitalizations in 2007 for each disease from the Health Care Utilization Project (HCUP) Inpatient Database, provided by the Agency for Healthcare Research and Quality (AHRQ). Hospital stays provide a measure of disease prevalence that incorporates the cost of diseases.

Appendix S1 discusses these two measures sources in more detail. We dropped two diseases (smallpox and anthrax), for which there were zero

deaths or hospitalizations over the period studied, resulting in a sample of 107 diseases (of which twelve are supercategories).

Empirical Analyses

We began by estimating simple bivariate regressions relating the natural log of funding for a disease in 2008 to the natural log of each of the disease burden measures. In this log-log specification, estimated coefficients can be treated as elasticities, that is, the percent increase in funding implied by a percent increase in a disease burden measure. We estimated these models both overall and separately after excluding the supercategories.

Next, we examined whether particular types of grants were more responsive to disease burden than others. Specifically, we estimated bivariate log-log regressions to examine the responsiveness of RFA versus other grants, clinical trials versus other grants, and center versus other grants.

Finally, we took advantage of the fact that the RCDC data are disaggregated to estimate separate elasticities for the twelve institutes (or centers) with the most funding in 2008: the National Cancer Institute (NCI), National Institute of Allergic and Infectious Diseases (NIAID), National Heart, Lung, and Blood Institute (NHLBI), National Institute of Diabetes and Digestive Diseases (NIDDK), National Institute of Neurological Diseases and Stroke (NINDS), National Center for Research Resources (NCRR), National Institute of Mental Health (NIMH), National Institute of Aging (NIA), National Institute of Drug Abuse (NIDA), National Institute of General Medical Sciences (NIGMS), National Institute of Child Health and Human Development (NICHD), and National Eye Institute (NEI). We estimated these models only over diseases funded by a particular institute. This was both practical (eliminating zeros in the dependent variable and allowing for logarithmic transformation) and reasonable, since it limited our observations to those diseases plausibly at risk and thus eligible for funding by an institute.

Results

Overall

Figure 1 shows scatterplots of the raw data and the fitted regression lines, as well as the main estimates and model statistics. (The markers refer to the RCDC ID number listed in table S1).

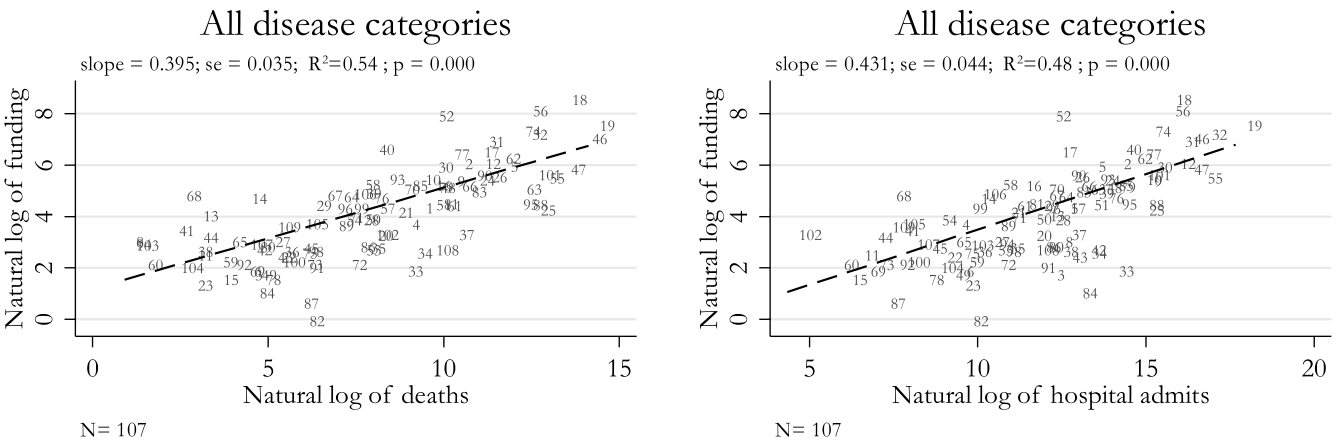

FIGURE 1. Funding in 2008 versus lagged deaths and hospital admits.

Notes: This figure shows scatterplots and fitted regression lines relating the natural log of funding for a disease in 2008 versus the natural log of deaths from that disease in 2005/2006 (left panel) and the natural log of hospitalizations in 2007 (right panel). The markers are the RCDC ID numbers for each disease listed in table S1. The bottom panels exclude "supercategories" (broad categories that span multiple other diseases in the RCDC). Slopes are elasticities; that is, they represent the percent increase in funding associated with a 1 percent increase in disease burden.
Source: Authors' analyses of data from the NIH's Research, Condition, and Disease Categorization database.

Continued.

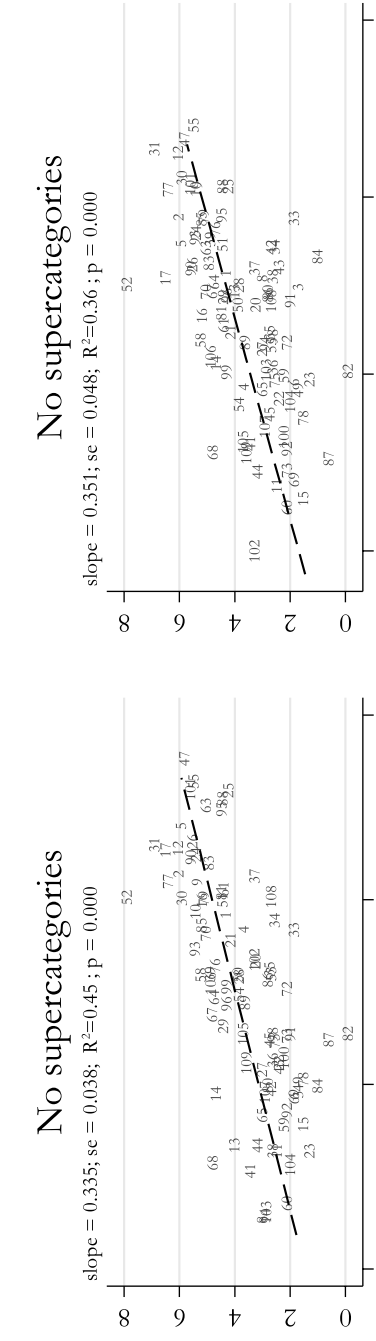

FIGURE 1. *Continued.*

The overall elasticity is .395, indicating that a 1 percent increase in deaths is associated with a .395 percent increase in funding. (In other words, a doubling of deaths implies about a 40 percent increase in funding.) Deaths explain 54 percent of the variation in funding. The estimated elasticity is lower (.34) after excluding supercategories. The elasticity for hospital admissions is higher than for deaths: .43 across all diseases and .35 excluding supercategories. Our basic analyses with the new RCDC data thus suggest a positive and statistically significant association between funding and both measures of disease burden, and each of the disease burden measures alone explains a substantial share of variation in NIH funding.

By Mechanism

Next we examined differences across types of NIH awards—RFAs, clinical trial funding, center funding, and contracts—to examine whether the more applied award types were more responsive. When examining each of these award types, we concentrated on those diseases with nonzero funding of that type. For example, when examining whether RFA or non-RFA funding was more responsive to disease burden, we limited the sample to diseases with at least one RFA.

Figure 2 shows the estimated elasticities and 95 percent confidence intervals. The top left panel shows the relationship between funding and deaths for all diseases, including supercategories. It shows that RFAs are slightly more responsive to deaths than non-RFAs; clinical trial funding is more responsive than nonclinical funding; and center funding is more responsive than noncenter funding. However, contract funding is less responsive to deaths than noncontract funding. The bottom left panels shows similar trends after excluding supercategories.

The right-hand side panels show similar patterns from using hospital admissions, rather than deaths, as the measure of disease burden. Targeted funding, particularly clinical and center funding, is more responsive to hospital admissions than untargeted funding. As was true for deaths, contract funding is less responsive to hospitalizations than noncontract funding.

Across the models, both the targeted and the untargeted funding measures have a positive and statistically significant (at the 5 percent level) association with disease burden, whether the burden is measured

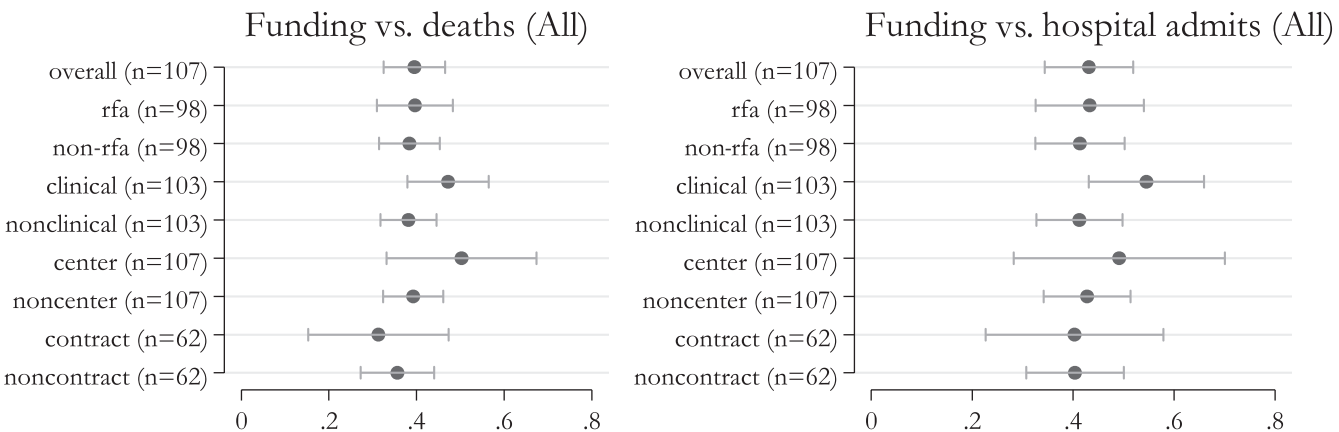

FIGURE 2. Estimated responsiveness of funding to deaths (left panels) and hospitalizations (right panels).

Notes: This figure plots coefficient estimates and 95 percent confidence intervals from bivariate regressions relating natural log of NIH funding (in 2008) to deaths in 2005/2006 (left panels) and hospitalizations in 2007 (right panels). The bottom panels show estimates from models excluding supercategories. For the RFA and non-RFA regressions, we include only diseases with at least one RFA grant and do the same for clinical and nonclinical regressions, center versus noncenter, and so forth. In each chart, "n" indicates the number of the 107 RCDC diseases over which the models were estimated. Estimates are elasticities; that is, they represent the percent increase in funding associated with a 1 percent increase in disease burden.

Source: Authors' analyses of data from the NIH's Research, Condition, and Disease Categorization database.

Continued.

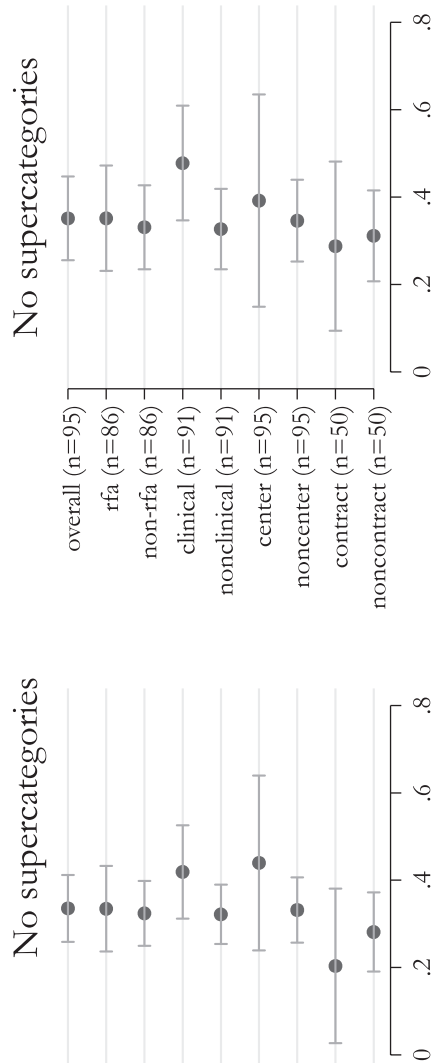

FIGURE 2. *Continued.*

by deaths or hospitalizations. Though the confidence intervals overlap, collectively the estimates provide some support for the argument that three of the targeted mechanisms (RFAs, clinical trial funding, and center grants) are relatively more responsive than untargeted funding, with the differences most pronounced for clinical trial funding and center grants.

By Institute

Finally, we used the data on disease specific funding by category to examine the relationships between funding and disease burden separately by institute or center, for the twelve largest funding institutes or centers. Figure 3 shows the estimates. The top left panel shows variation across the entities in responsiveness to deaths. Several of the institutes have elasticities greater than .4 (NCI, NCRR, NIA), while five of them have elasticities less than .2 (NIAID, NINDS, NIMH, NIDA, NEI). Indeed, for four institutes (NIAID, NIMH, NIDA, and NEI), the 95 percent confidence intervals include zero; that is, the relationship between funding and deaths is not statistically significant at the 5 percent level. The highest responsiveness is at NCRR, a heavy funder of clinical and translational research. (NCRR was recently replaced by the National Center for Advancing Translational Sciences; see Kaiser 2010.)

The general trends (in responsiveness of funding to deaths) across institutes are similar after excluding supercategories, as the bottom left panel shows. As in figure 1, the institute-specific estimates of responsiveness to deaths are generally lower after excluding supercategories.

The right-hand side panels show the analogous results for hospitalizations. With this burden measure, too, there is variation, and the trend in responsiveness across institutes is similar to that already observed for deaths.

Discussion

Previous work, based on a smaller set of diseases, provides mixed evidence on the responsiveness of NIH funding to disease burden, including both deaths and hospitalizations, and to other measures of disease burden. Our study used data on 107 diseases categorized by the RCDC, as well as the funding numbers officially endorsed by the NIH. We found strong

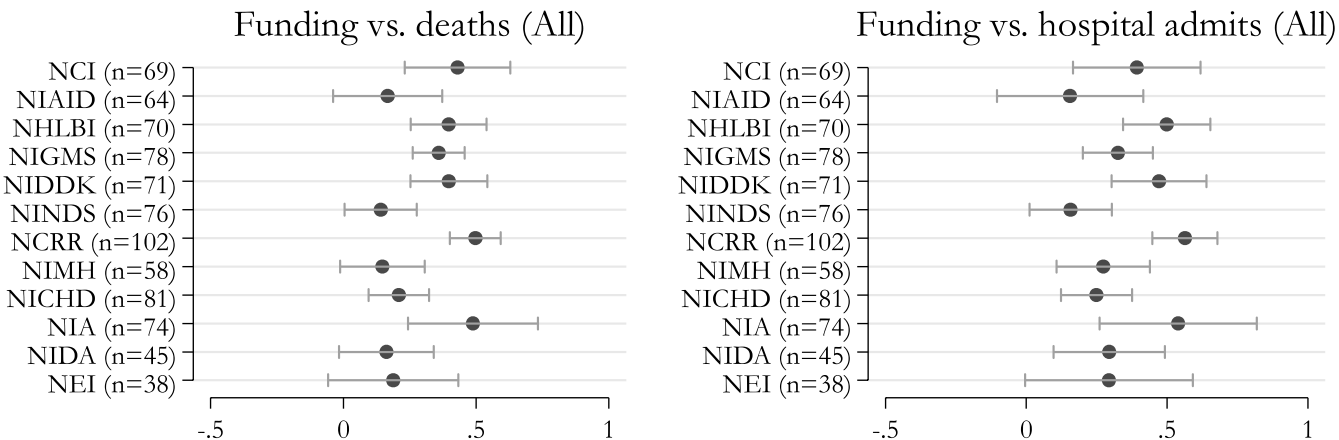

FIGURE 3. Estimated responsiveness of funding to deaths (left panels) and hospitalizations (right panels) for the twelve largest funding institutes and/or centers.

Notes: This figure plots coefficient estimates and 95 percent confidence intervals from bivariate regressions relating natural log of NIH funding (in 2008) to deaths in 2005/2006 (left panels) and hospitalizations in 2007 (right panels), estimated separately by institute. The models were estimated for the twelve largest institutes, ranked by total 2008 funding. The bottom panels show estimates from models excluding supercategories. In each chart, "n" indicates the number of the 107 RCDC diseases in our sample funded by the institute, that is, the sample size for that institute's regression. Slopes are elasticities; that is, they represent the percent increase in funding associated with a 1 percent increase in disease burden.

Source: Authors' analyses of data from the NIH's Research, Condition, and Disease Categorization database.

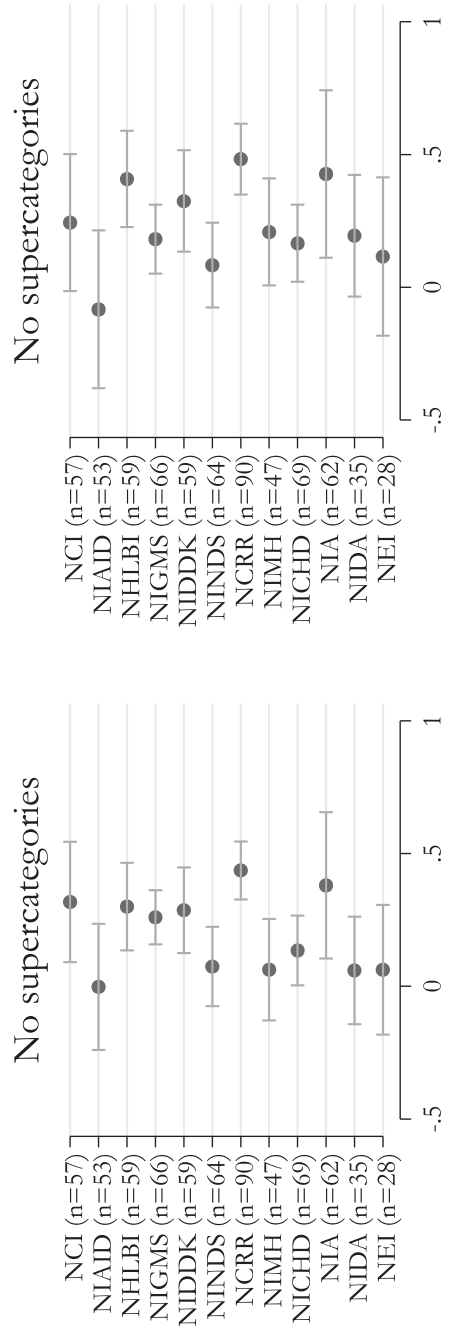

FIGURE 3. *Continued.*

and statistically significant associations between NIH funding and each measure of disease burden.

We also found overall elasticities of funding with respect to deaths to be .4 ($p < .01$) and funding with respect to hospitalizations to .43 ($p < .01$). In other words, a doubling of hospitalizations was associated with a 43 percent increase in funding, and a doubling of deaths, with a 40 percent increase. By comparison, using pre-RCDC funding data for twenty-nine diseases, Gillum and colleagues (2011) estimated elasticities of funding with respect to hospitalizations to be .21 ($p = .05$), and funding with respect to mortality to be .12 ($p = .08$). Previous work on the same twenty-nine diseases (Gross, Anderson, and Powe 1999) did not report elasticities but found a statistically significant relationship between funding and deaths ($p = .03$). Although these researchers did not collect data on the number of hospitalizations, they did examine hospital days and detected no significant relationship between this variable and funding ($p = .21$).

In addition to being officially endorsed by the NIH, a second benefit of the RCDC data is that they facilitate an examination of channels of potential responsiveness. Earlier scholarship did not devote much attention to these channels of possible influence, despite statements from various observers that the NIH allocation process—dominated by investigator-initiated grants and scientific merit—may not be well positioned to set priorities according to health burden.

Across the different models, we found that compared with other types of grants, RFA-funded research, clinical trials, and center grants each are more responsive to both deaths and hospitalization. (However, both contract and noncontract grants are similarly responsive to disease burden.) While the confidence intervals typically overlap, the overall patterns in the point estimates are consistent with the argument that it is more feasible for the agency to respond to disease burden considerations through these three "directed" or "applied" grants than through investigator-initiated "basic" research.

We also showed differences across the major funding institutes in their responsiveness to disease burden. (The estimated elasticities are statistically significant for many, but not all, of the institutes.) This is consistent with previous arguments that the priority-setting processes at the different institutes and centers are quite different.

Together with other work (Sampat 2012), the results point to an interesting political economy at the NIH. Congress and disease

interest groups may have an important role in emphasizing particular broad diseases, resulting in more funding for certain institutes than others. And the more directed grant mechanisms may serve as important "safety valves" in the allocation process to respond to pressure to address particular health priorities (and/or political considerations) in a more fine-grained way than is possible through investigator-initiated "basic" research grants. From this perspective, Congress and the more directed grant mechanisms would serve an important role in promoting the NIH's responsiveness to health objectives.

The results also suggest, however, that even the less directed (non-RFA, noncenter, nonclinical, noncontract) grants show positive and statistically significant responsiveness to both measures of disease burden. Moreover, there is evidence of responsiveness to disease burden even within the majority of the institutes. This is perhaps even more surprising, given the concerns noted earlier that the peer review process is narrowly focused on scientific considerations. These findings may indicate that there are levers to respond to health priorities even in peer reviews of investigator-initiated grant applications. Perhaps external peer reviewers focused on scientific merit as well as disease burden. Similarly, scientists' research choices (and thus "untargeted" investigator-initiated proposals) may themselves be driven by not only scientific but also demand-side considerations. Another possibility is that the second stage of the NIH's dual peer-review process—in which the institutes' advisory councils emphasize both science and relevance to the institutes' "goals and needs" (NIH 2012)—is generating the estimated responsiveness. Still another mechanism, more controversial, is "relabeling": the content of research may not be targeted and may be quite fundamental, although the RCDC labels assigned to it could be more targeted to specific disease categories.

Limitations

There are several limitations on what we can say about this analysis. The first is that the estimates are based on cross-sectional, not longitudinal, variation, since the RCDC funding data became available only recently and cover a short time span. The second limitation is that the results are based on RCDC-listed diseases only, which may not be a representative subset of all diseases. The RCDC categorization builds on the existing infrastructure of "diseases of interest," which were preserved "to enable

NIH to respond to requests from Congress or others" (IOM 1998, 21). If the diseases listed on the RCDC are those for which NIH responsiveness is greater, for example, estimates based on this sample would be higher than those for non-RCDC-listed diseases.

More generally, while the RCDC has advantages over previous data (including transparency, links to a large number of diseases, and consistency in how definitions are applied across grants from different institutes and centers), it also has real limits. One is that the categories are neither complete nor mutually exclusive (as was true of earlier funding data provided by the agency). Even more fundamental is the question of whether any parsing of funds by disease is possible for an agency that funds "basic" research, and in a context in which research in one disease area sometimes (perhaps often) serendipitously leads to progress in another.

Another issue is that factors beyond disease burden likely matter for the NIH's funding choices as well (IOM 1998; Varmus 1999b). In particular, there is broad recognition that scientific opportunity (the expected success of particular lines of research) also ought to factor into funding choices, even if the agency were interested only in maximizing health benefits from its research (Lichtenberg 2001; Zeckhauser 1967). In other words, the right question may be how responsive the NIH is *conditional* on scientific opportunity. Unfortunately, scientific opportunity is difficult to operationalize in large samples, so our analysis does not speak to this broader question.

This analysis examines only two measures of U.S. disease burden, those for which data were systematically available for all diseases on the RCDC list. The correct measure to use is still highly contested (Gold and Fryback 2002), and it is possible that our results would be different if we had examined other measures, including those that account for morbidity (e.g., DALYs). For example, several of the institutes for which we estimated low elasticities with respect to deaths and hospitalizations may instead be responsive to other measures of the U.S. burden of disease. Moreover, the disease burden outside the United States (however measured) could also affect allocation choices. Previous analyses (Gross, Anderson, and Powe 1999) found different results based on different measures of burden. Extending our analysis to examine other measures is an important task for future research.

While our analyses examined only extramural funding, another potentially interesting safety valve mechanism is the intramural program.

Though a small share of the agency's budget (about 10 percent in 2008), intramural research has historically been important in helping the agency respond to public health emergencies, including HIV-AIDS (Hardin 2012). Priority setting, including responsiveness to disease burden, would seem to be more viable through intramural funding if such funding were easier to steer. It would be interesting, in future research, to examine this formally.

We close by emphasizing that the broad set of potential desiderata and considerations complicates any analysis of priority setting at the NIH (IOM 1998), making the application of standard approaches in health care priority setting highly challenging (Varmus 1999b). While the RCDC was a significant step in making NIH decision making more transparent and accountable, more specific information on what criteria the agency considers in its allocation decisions, where in the process it does so, and the weights it places on differing priorities and goals would help facilitate future evaluation and might be even more important than better data.

References

Bowker, G. 1996. The History of Information Infrastructures: The Case of the International Classification of Diseases. *Information Processing and Management* 32:49–61.

Bowker, G., and L. Starr. 1999. *Sorting Things Out: Classification and Its Consequences.* Cambridge, MA: MIT Press.

Callahan, D. 1999. Shaping Biomedical Research Priorities: The Case of the National Institutes of Health. *Health Care Analytics* 7:115–29.

Cohen, J. 1993. Conflicting Agendas Shape NIH. *Science* 261:1674–79.

Fuchs, V.R. 1975. *Who Shall Live? Health, Economics, and Social Choice*. New York: Basic Books.

Gillum, L.A., C. Gouveia, E.R. Dorsey, M. Pletcher, C.D. Mathers, C.E. McCulloch, and S.C. Johnston. 2011. NIH Disease Funding Levels and Burden of Disease. *PloS One* 6:e16837.

Gold, M., and S.D. Fryback. 2002. HALYs and QALYs and DALYs, Oh My: Similarities and Differences in Summary Measures of Population Health. *Annual Review of Public Health* 23:115–34.

Gross, C., G. Anderson, and N. Powe. 1999. The Relation between Funding by the National Institutes of Health and the Burden of Disease. *New England Journal of Medicine* 340:1881–87.

Hardin, V. 2012. *AIDS at 30: A History*. Washington, DC: Potomac Books.

IOM (Institute of Medicine), Board on Health Sciences Policy. 2004. *NIH Extramural Centers Program: Criteria for Initiation and Evaluation*. Washington, DC: National Academies Press.

IOM (Institute of Medicine), Committee on the NIH Research Priority-Setting Process. 1998. *Scientific Opportunities and Public Needs: Improving Priority Setting and Public Input at the National Institutes of Health*. Washington, DC: National Academies Press.

Johnson, J. 1998. Disease Funding and NIH Priority Setting. CRS Report for Congress. Available at http://www.cnie.org/NLE/CRS reports/science/st-35.cfm (accessed December 10, 2011).

Kaiser, J. 2010. A Government Niche for Translational Medicine and Drug Development. *Science* 330:1462–63.

Kastor, J. 2010. *The National Institutes of Health 1991–2008*. New York: Oxford University Press.

Lichtenberg, F. 2001. The Allocation of Publicly Funded Biomedical Research. In *Medical Care Output and Productivity*: *Studies in Income and Wealth, Vol. 62*, ed. D. Cutler and E. Berndt, 565–89. Chicago: University of Chicago Press.

Mowery, D., R. Nelson, and B. Martin. 2010. Technology Policy and Global Warming: Why New Policy Models Are Needed (or Why Putting New Wine in Old Bottles Won't Work). *Research Policy* 39(8):1011–23.

Mushkin, S. 1979. *Biomedical Research: Costs and Benefits*. Pensacola, FL: Ballinger.

National Research Council. 2003. *Enhancing the Vitality of the National Institutes of Health: Organizational Change to Meet New Challenges*. Washington, DC: National Academies Press.

NIH (National Institutes of Health). 1997. *Setting Research Priorities at the National Institutes of Health*. Washington, DC.

NIH (National Institutes of Health). 2011a. *Estimates of Funding for Various Research, Condition, and Disease Categories (RCDC)*. Available at http://report.nih.gov/categorical_spending.aspx (accessed May 2, 2011).

NIH (National Institutes of Health). 2011b. *Research Portfolio Online Reporting Tools (RePORTER)*. Available at http://projectreporter.nih.gov/reporter.cfm (accessed March 15, 2011).

NIH (National Institutes of Health). 2012. *Peer Review Process*. Available at http://grants.nih.gov/grants/peer_review_process.htm (accessed November 3, 2012).

Rettig, R.A. 1977. *Cancer Crusade: The Story of the National Cancer Act of 1971*. Washington, DC: Joseph Henry Press.

Sampat, B. 2012. Mission-Oriented Biomedical Research at the NIH. *Research Policy* 41(10):1729–41.

Strickland, S.P. 1972. *Politics, Science, and Dread Disease: A Short History of United States Medical Research Policy*. Cambridge, MA: Harvard University Press.

Varmus, H.E. 1997. *Testimony on Setting Research Priorities at the NIH*. House Committee on Appropriations, Subcommittee on Labor, Department of Health and Human Services, Education, June 10, 1997.

Varmus, H.E. 1999a. Evaluating the Burden of Disease and Spending the Research Dollars of the National Institutes of Health. *New England Journal of Medicine* 340:1914–15.

Varmus, H.E. 1999b. *Statement on Funding Allocation for Disease Research*. Senate Committee on Appropriations, Subcommittee on Labor, Department of Health and Human Services, Education, May 6, 1999.

Weisbrod, B.A. 1983. *Economics and Medical Research*. Cambridge, MA: AEI Press.

Zeckhauser, R. 1967. *Some Thoughts on the Allocation of Resources to Bio-Medical Research*. Washington, DC: U.S. Department of Health, Education and Welfare, Office of the Assistant for Planning and Evaluation.

Acknowledgments: This research was supported by a Robert Wood Johnson Foundation Investigator Award in Health Policy Research to Sampat. We thank Larry Brown, Sherry Glied, Tal Gross, Matthew Neidell, Sejal Patel, Michael Sparer, and Jennifer Washburn for comments and suggestions on a previous draft. Parts of this manuscript were written while Sampat was a visiting scholar at the Hastings Center for Bioethics.

Supplementary Material

Additional supporting information may be found in the online version of this article at http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1468-0009:

Appendix S1: discusses coding rules in detail.
Table S1: shows the RCDC category, funding, and ICD codes for these RCDC categories.